Berkeley Lab Cluster Update

Bill Saphir
wcsaphir@lbl.gov
3/9/00

What do these have in common?

- Bach's *Well Tempered Klavier*
- Pepper
- Supercomputers

Machines

Production Clusters (?)

- PDSF
  - Serves HEP and NP communities
  - 70 nodes/124 cpus/5TB heterogeneous. Doubling FY00
  - 24x7; high utilization; high availability; venture capital
- QDSF
  - General purpose
  - >60 Gflop/s peak
  - FY00

Development

- x86 testbed
  - 36 nodes
  - FE/Giganet
- Alpha testbed (“Babel”)
  - 12 nodes (DS10)
  - Servernet II/Myrinet/Gigabit

Machines 2

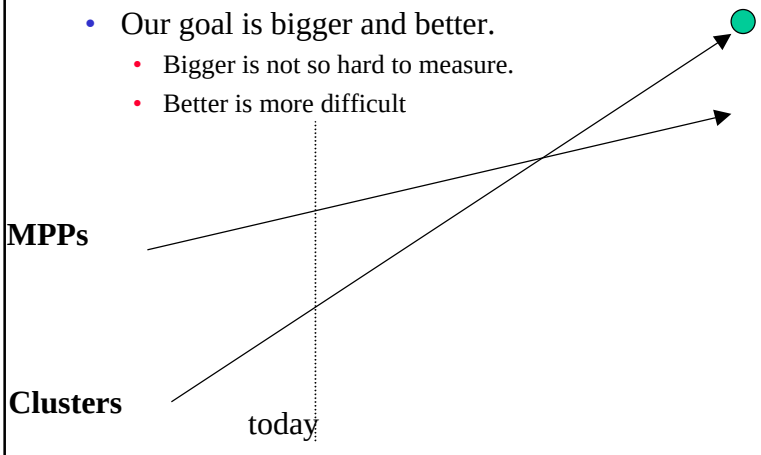
- IBM SP
  - 2048+ processors this year
  - > 3 Tflop/s peak
  - Global parallel filesystem for home directories
  - Many MPI-2 features
  - O($10k)/processor
- T3E/900
  - 640 processors
  - 95% utilization
  - Checkpoint/restart
  - 2us latency for those who need it

R&D/Software

- BLD
  - Plug and play cluster software
  - Administration tools
  - Integration
- M-VIA
- MVICH
- Channel bonding
- Measurement

Measuring progress

- How does a Linux cluster compare to vendor systems? SP; T3E; Red; O2K?
- Our goal is bigger and better.
  - Bigger is not so hard to measure.
  - Better is more difficult



Metrics

- Power consumption, cost
- % availability
- % utilization
- MTBI
- MTTR
- Performance on app-level benchmarks.
- Actual performance delivered (counters)
- Effective System Performance test
- Amount of administration for certain level of service
- Quality of user environment?
- What about difficult scientific problems?
- Ask the users!

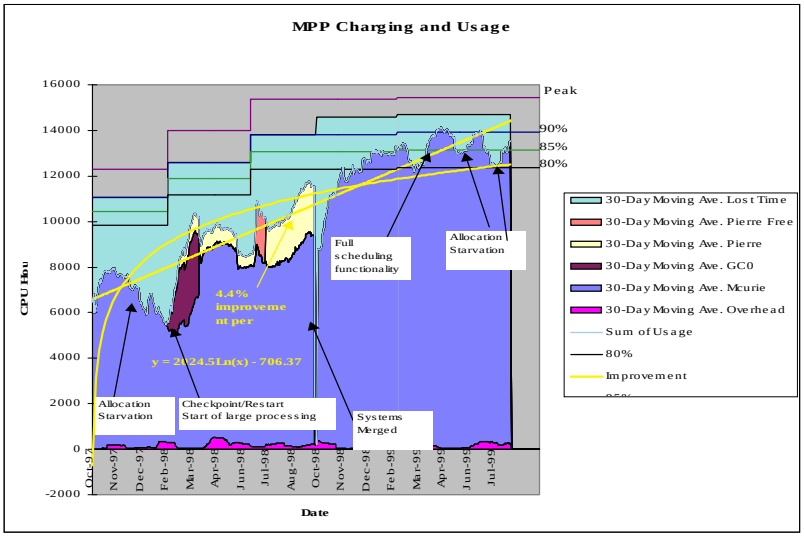
System Metrics

NERSC System Metrics for FY99

Measured (Goal)

	% Availability		MTBI	MTTR
Systems	Scheduled	Overall	(Hours)	(Hours)
Vector Systems	99.58 (96	99.04 (96)	361(96)	3.2(4.0)
Storage Systems	99.39 (96)	98.63 (95)	169(96)	2.8(4.0)
Parallel Systems	97.83 (96)	96.02(95)	81(96)	3.1(4.0)
Workstations				
Servers (fs/gw)	100(96)	100 (96)	NA(340)	NA(8.0)
Clusters	99.7(96)	99.5(95)	532(40)	1.7(8.0)

Utilization



How to be successful

- Measure measure measure -- Open metrics
- Release release release -- Open source
- R&D&D&D
- Avoid ad hoc solutions. design; integrate; document. Build for tomorrow, not today.